AD-A259 397

15010

(2)

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

DTIC ELECTE DEC 3 0 1992 S A D

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 92 0931
6a. NAME OF PERFORMING ORGANIZATION Department of Psychology SUNY Stony Brook	6b. OFFICE SYMBOL (If applicable)
7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research	
6c. ADDRESS (City, State and ZIP Code) Stony Brook, NY 11794-2500	7b. ADDRESS (City, State and ZIP Code) Bolling Air Force Base, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR/NL	8b. OFFICE SYMBOL (If applicable) NL
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-91-0378	
8c. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
PE 61102F	PR 2313	TA A6	

11. TITLE (Include Security Classification) Signal- and Listener-based Factors in Complex Auditory Pattern Perception

12. PERSONAL AUTHOR(S)
Arthur G. Samuel

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Annual Technical	FROM 9/15/91 TO 9/14/92	1992-10-25	40

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The research conducted during the first year of funding addressed fundamental issues in the perception and internal representation of speech. Three separate lines of research were designed to clarify the relationship between low-level (phonemic) speech codes and higher-level (lexical) ones. Two of the approaches involved tests of whether the activation of a lexical (word) representation increases the activation of its components (i.e., the phonemes that compose it). One of these lines investigated the extent to which listeners perceptually restore deleted phonemes, while the other examined how rapidly listeners can detect a pre-specified phoneme. The third line of research tested whether the existence of one lexical representation (e.g., the word "dent") inhibits the perception of a similar word (e.g., "tent"). Together, the various investigations are intended to clarify the structure and processes of the human speech perceptual system.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Dr. John Tangney	202-767-5021	AFOSR/NL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

6 NOV 1992

I. Summary

The research conducted during the first year of funding addressed fundamental issues in the perception and internal representation of speech. Three separate lines of research were designed to clarify the relationship between low-level (phonemic) speech codes and higher-level (lexical) ones. Two of the approaches involved tests of whether the activation of a lexical (word) representation increases the activation of its components (i.e., the phonemes that compose it). One of these lines investigated the extent to which listeners perceptually restore deleted phonemes, while the other examined how rapidly listeners can detect a pre-specified phoneme. The third line of research tested whether the existence of one lexical representation (e.g., the word "dent") inhibits the perception of a similar word (e.g., "tent"). Together, the various investigations are intended to clarify the structure and processes of the human speech perceptual system.

II. Research Objectives

The objective of the research project is to delineate principles that underlie the perception of complex auditory patterns. The stimuli used are speech and musical patterns of varying complexity. A wide array of experimental procedures and analyses are used to try to determine properties that are true of the perception of complex auditory patterns across stimulus domains. In addition, we also are interested in discovering any principles that are domain specific (e.g., as "categorical perception" has traditionally been claimed to be a principle of perception specific to the speech domain). The various experimental investigations in the project may be broadly grouped into studies of signal-based factors, and studies of listener-based factors. The former group includes experiments that explore how properties of the input signal determine perception, while the latter group includes studies of how listeners' expectations influence perception/performance. The former group primarily focusses on early representations of the signal, and the latter includes higher-level factors (including, but not limited to, attentional influences). The long-term goal of the research is to understand both signal-based and listener-based factors, and their interaction in the perception of complex auditory patterns.

III. Progress Toward Research Objectives

The first year of funding for AFOSR 91-0378 has coincided with our first year here at Stony Brook. During this year, we have conducted our research while establishing our new laboratory. The three lines of research that we pursued nicely illustrate our transition year: One line was conducted at a collaborator's laboratory (Mark Pitt, at Ohio State), one was conducted using our old lab's equipment in a borrowed lab here, and the third was conducted using our new facilities in our new laboratory. At this point, the new laboratory is quite functional, which will allow us to make substantial progress in the coming year.

All three ongoing projects involve investigations of the interaction of relatively low-level representations of speech signals (phonemic codes) with higher-level representations (lexical codes). One research program has investigated the extent to which activation of a lexical code increases the activation of its component phonemes, indexing this by listeners' ratings of how much they perceptually restored phonemes that were deleted from the signal. A second line of research pursues similar theoretical questions, but using detection times for specified phonemes as the dependent measure. The third research effort tests whether the existence of one lexical representation (e.g., of the word "dent") inhibits the perception of a similar word (e.g., "tent"). Collectively, these research programs are clarifying how different levels of analysis of the complex speech signal interact during perception.

A. Perceptual Restoration. The perceptual system is designed to produce a filtered version of reality: Incomplete or ambiguous stimuli will usually be perceived as more complete and less ambiguous than the input. The operation of such restoration processes makes it clear that a full understanding of perception must extend beyond the specification of signal-based factors. Studies of perceptual restoration effects, beginning with Warren's (1970) seminal paper, have begun to clarify the perceptual architecture. The discrimination methodology introduced by Samuel (1981) has proved very useful in distinguishing between perceptual restoration and post-perceptual biases; these results have played an important role in the debate over modular versus interactive architectures (cf Fodor, 1983).

In the discrimination paradigm, two types of stimuli are constructed. In one type, a segment of the speech is digitally removed, and replaced by white noise of approximately the same amplitude. In the second stimulus type, the noise is superimposed on the corresponding segment. Listeners must judge whether a given stimulus is intact (i.e., has noise superimposed on a speech segment), or not (i.e., has noise replacing a speech segment). Of course, if the perceptual system restores deleted speech segments, then the two stimulus types will be difficult to discriminate because both will sound intact, with superimposed noise.

Insert Figure 1 Around Here

The discrimination methodology relies on signal detection analyses to separate perceptual effects (d') from postperceptual biases (Beta). Earlier studies of perceptual restoration have assumed an underlying signal detection model of the sort shown in Figure 1. In this model, the assumption is that when listeners hear a stimulus (e.g., a word with some noise in it), there is some internal representation that includes some metric of how intact the stimulus sounded. More specifically, when stimuli have a noise burst replacing a phoneme, relatively low intactness values are generated; the usual signal detection distribution assumptions are made. For stimuli in which noise is simply added to the signal (without replacing a phoneme), we assume an intactness distribution that is higher. The model is shown in the top panel of Figure 1. Previous restoration studies looked for lexical influences on perception of phonemes by increasing some lexical factor (e.g., priming a word, or contrasting real words [lexical] with pseudowords), and looking for a reduction in the ability to discriminate a truly intact stimulus ("added") from one that was perceptually reconstructed ("replaced"). As shown in the bottom of Figure 1, the reduced d' is predicted if the lexical information is used by the perceptual process to fill in the replaced phoneme, thus making the "replaced" stimulus sound more intact.

Insert Figure 2 Around Here

Our new research explores a possibility illustrated in Figure 2. If the lexical information is used to both restore "replaced" phonemes and to increase the perceived intactness of "added" ones (which though technically intact are nonetheless degraded by the noise), then both distributions could shift up. This possibility is illustrated in the bottom half of Figure 2. Note that d' does not change from the initial case (e.g., unprimed words) to the second situation (e.g., primed words). If this really happens, then we could incorrectly infer that the lexical information was not being used by the perceptual process.

The work we have done this year is an attempt to determine what actually happens to the distributions, and to try to account for varying results with different stimulus sets. The basic procedure is to present listeners with an "added" or "replaced" stimulus, followed by the stimulus without any noise addition or replacement. The second item serves as a standard, and subjects are told to compare the quality of the first item to the standard. They rate the first item on an 8-point scale, with 1 indicating that the test item was highly degraded relative to the standard, and 8 meaning that the test item was just as good as the standard. Subjects are not told that there are two different noise manipulations, but we partition the responses on the basis of noise addition versus replacement. This allows us to recover an estimate of the hypothesized underlying intactness distributions for each stimulus type.

This year we have used this new method in three experiments. In two experiments we used a set of 90 words and 90 pseudowords; these were the same words and pseudowords used in the original discrimination study (Samuel, 1981). In one of these experiments, we

compared the ratings for words to those for pseudowords. In the second, we compared ratings for primed versus unprimed words, using the same 90 words. The third rating experiment used a different stimulus set (108 words, 108 pseudowords), one that has given different results when these stimuli were used in the standard paradigm.

In our first rating experiment, each trial consisted of a test item (added or replaced), followed by the intact (noise-free) version of the test item, with an ISI of 750 ms. The second item actually consisted of a leading noise burst, the intact standard, and a following noise burst. The bracketing noise bursts were included because in pilot testing, we found that there was a tendency for the ratings to be influenced by the loudness of the noise burst in the test item (the noise amplitude was based on the amplitude of the critical phoneme, and therefore varied across stimuli). To negate this tendency, noise bursts were put around the intact standard that matched the amplitude and duration of the noise in the test token. This was accomplished by using a copy of the noise segment actually used in the "replaced" version of the stimulus. The net result was that there was no longer a tendency for the noise amplitude to influence the ratings, because for every trial, the noise in the test token matched noises surrounding the standard. The results were scored by compiling a distribution of ratings for each condition. In addition, mean ratings in each condition for each subject were computed. These were used in a set of analyses of variance.

Insert Figure 3 Around Here

Figure 3 presents an overall summary of the ratings, broken down by lexical status (word versus pseudoword) and intactness (added versus replaced). Inspection of the Figure leads to three immediate impressions: (1) Overall, "added" stimuli were rated as more intact than "replaced" stimuli; ratings for the former peak around a value of 7, versus a peak around 5 for the latter. (2) The word and pseudoword distributions are very similar for "added" stimuli, with a slightly higher distribution for words than pseudowords. (3) The word and pseudoword distributions for "replaced" stimuli are shifted relative to one another, with words yielding higher intactness ratings than pseudowords. These results are very much in accord with the original logic described by Samuel (1981), and schematized in Figure 1. In most of our subsequent analyses, we have concentrated on the intactness ratings for "replaced" stimuli, since these are the ones that usually show effects; important exceptions are described below.

Insert Figure 4 Around Here

Figure 4 presents the intactness ratings for the "replaced" stimuli, broken down by lexicality and critical phoneme position. Note that there is no lexicality effect in initial position, there is a moderate effect medially, and there is a very large shift in final position:

In final position, the distribution for pseudowords peaks around 3 or 4, while for words the peak is around 6. The results in final position in particular are remarkably similar to the assumptions underlying the original technique.

Insert Figures 5, 6, and 7 Around Here

Figures 5, 6, and 7 present the final breakdown of the first experiment's results, in terms of the phone class of the probed location. It may simplify matters to note that the results are very similar for stops and fricatives (Figure 5), and for the vowels and liquids (Figure 7). Thus, conceptually we may deal with three cases: stops/fricatives, nasals, and vowels/liquids. For the first type, the results are very simple: There is virtually no difference between words and pseudowords, or for that matter, between "added" and "replaced". As demonstrated in several studies, these stimuli are discriminated at near-chance levels.

The results for nasals are entirely different. As is clear in Figure 6, for both "added" and "replaced" stimuli, the word distributions are noticeably higher than those for pseudowords. This is exactly the kind of "double perceptual" shift shown in Figure 2, a type of shift that will not produce reliable changes in d'. Recall that this kind of effect was hypothesized to be possible if the experimental manipulation shifted the perceived intactness of both the truly disrupted "replaced" stimuli and the merely degraded "added" ones.

If nasals, fricatives, and stops all produce results that will not yield observable d' effects in the standard paradigm, then where do the d' effects come from? Figure 7 suggests that most of the effect must derive from liquids and vowels, the two classes that paradoxically are the most weakly restored (overall ratings of 4.70 and 5.02 for vowels and liquids, versus 5.24, 5.79, and 6.00 for nasals, fricatives, and stops). As shown in Figure 7, liquids show no lexical effect at all for "added" stimuli, and vowels show little, while both classes reveal systematic shifts for the "replaced" versions. This is of course the pattern that will produce observable d' effects. Thus, collectively, the "replaced" distribution will shift up in words, relative to pseudowords, due to changes found with nasals, liquids, and vowels, while the "added" distribution will change little because four of the five phone classes do not show any shifts for the "added" case.

The first use of the new rating procedure proved to be extraordinarily useful. The results provided a visible estimate of the hypothetical distributions underlying Samuel's (1981) application of signal detection theory. These estimated distributions confirmed that lexical status significantly affects perceptual restoration, with the hypothesized shift in the distributions occurring primarily in later positions. In addition, the first experiment provided a strong clue to the reason behind the failure to find lexical effects in several earlier experiments: The stimuli used in those experiments had critical phonemes chosen from the

fricative, stop, and nasal phone classes. Note that the analyses just presented indicate that these are precisely the least likely stimuli to produce observable d' changes.

Before examining this issue directly in Experiment 3, we will first briefly present an exploration of the effect of priming. In Samuel's (1981) original study, priming a word significantly reduced subjects' ability to distinguish "added" from "replaced" versions of words. This loss was localized, with the effect occurring in final position. The same pattern was found in both instantiations of the priming experiment that were run, with the two varying in the prime-test spacing (200 msec ISI in one case, and approximately 800 msec in the other). Several later experiments, using a different set of stimuli, failed to find such priming effects.

Our second rating study used the same 90 words as Experiment 1, with priming procedures. This provides estimates of the distributions for "added" and "replaced" words, in order to clarify the effect of priming. In the primed condition, the test word was preceded by the prime, with an ISI of 500 msec. The prime did not include the bracketing noise bursts that were used with the standard. Thus, a trial consisted of the prime followed by the test word followed by the noise-bracketed standard. The subjects in Experiment 2 also did an unprimed word condition, identical to the word condition of the previous experiment.

Insert Figure 8 Around Here

The rating scores for each condition were averaged and analyzed as in the first experiment. Figure 8 presents the primed and unprimed rating distributions, collapsed across the five phone classes. The original priming result (Samuel, 1981) was that priming should lower d' scores by bringing the "replaced" distribution closer to the "added" distribution. In the present experiment, this translates to an interaction between priming and stimulus type. This critical interaction was reliable, reflecting the fact that in the unprimed case the average "added" rating (6.05) was further from the average "replaced" value (4.79) than the primed "added" mean (6.11) was from the primed "replaced" average (5.03). Thus, the rating procedure indicates that the shift in the distributions follows the predicted pattern. However, although the interaction was reliable and as predicted, inspection of Figure 8 makes it clear that there was a more pronounced "double shift" than the "single shift" (of the "replaced" distribution) reflected in the interaction. The fact that priming tends to shift both distributions means that finding d' changes in the standard paradigm may be difficult: The presence of such effects must be due to cases in which the "double shift" does not swamp the effect of the "replaced" distribution shifting differentially. Once again, the rating procedure has provided theoretically useful estimates of the underlying distributions.

As noted previously, several studies run after Samuel's (1981) paper failed to find d' differences between words and pseudowords. In those experiments, a new 108-word stimulus set was constructed to use as the basis for 216 test words and 216 test pseudowords.

The design of the set was guided by a desire to create a testing situation with relatively high uncertainty, and relatively difficult discrimination conditions (to keep performance away from ceiling). This led to a stimulus set with three word lengths, three critical phoneme positions, and three phone classes. As noted, this stimulus set failed to produce a pseudoword advantage, despite several methodological changes intended to control for possible artifacts.

The results of the first rating experiment provided an indication of a possible source of this failure. Recall that in the rating task, the observed distributions for stops and fricatives did not reflect any lexical effect: For both words and pseudowords, the "added" and "replaced" functions were very similar, reflecting the high level of perceived intactness for all of these stimuli. For nasal critical phonemes, a very different pattern was found, but one that also would not produce any observable d' effect: Both the "added" and the "replaced" intactness ratings were higher for words than for pseudowords, the "double shift" that leaves the distance between the distribution's means unchanged.

Because the critical phonemes for the stimuli in the studies that failed to find d' effects were stops, fricatives, and nasals, it seems reasonable to speculate that the failure to find a pseudoword d' advantage was due to the nature of the critical phonemes. We have run a third experiment to test this speculation, using the rating paradigm to obtain estimates of the underlying intactness distributions for the stimuli used in those experiments. The third experiment was procedurally identical to the first, but used the 108-item stimulus set that had failed to produce lexical effects in the standard paradigm.

Insert Figure 9 Around Here

The results were scored and analyzed as in the first experiment. The distributions for stops, fricatives, and nasals are shown in Figure 9. In a lexicality X stimulus type X phone class ANOVA, all three main effects were reliable. Words (5.97) yielded higher ratings than pseudowords, added stimuli (6.08) were rated as more intact than replaced ones (5.58), and the nasals (5.59) received lower ratings than fricatives (5.81) or stops (6.09). There was also a robust interaction of stimulus type and phone class. This interaction was due to the fact that the "added" distributions are quite similar for the three phone classes (see top panels of Figure 9), but the "replaced" distributions for nasals are much lower than the comparable results for stops and fricatives (bottom panels). In fact, as is clear in the Figure, the "replaced" functions for stops and fricatives are virtually identical to the corresponding "added" functions. This replicates the stop and fricative results found in Experiment 1 with the 90-item stimulus set, and confirms the difficulty of finding significant changes in d' using the original discrimination paradigm with these stimuli.

The functions for the nasal stimuli also reveal the expected difficulty, although not quite as clearly as in the first experiment. As in that study, <u>both</u> the "added" and the "replaced" distributions for words are shifted up relative to those for pseudowords: In the

"added" case, the average rating for words (6.16) was 0.4 units higher than the average rating for pseudowords (5.86). For "replaced" stimuli, comparable difference of .36 units was based on a word mean of 5.30, versus a mean of 4.94 for pseudowords. This double shift is just what was found for the nasals in Experiment 1. The shifts with the present stimuli are a bit smaller than in that experiment. A plausible basis for this is that in the current experiment, two thirds of the stimuli were virtually impossible to work with (the stops and fricatives), and subjects may have developed less optimal judging procedures under these circumstances. In any event, the phone class analysis clearly confirms the hypothesis that stops and fricatives produce what are essentially stimulus-bound ceiling effects, and nasals yield the double shift pattern that was hypothesized to be expected if a manipulation affected the perception of both "replaced" and degraded stimuli.

Taken together, the three rating experiments have provided a remarkably useful insight into how listeners perceive incomplete signals. The results have validated Samuel's (1981) original logic and techniques, and have shown when the original paradigm will fail. We are continuing this line of investigation.

B. A Test for Lexical Inhibition. The restoration work just discussed tests one prediction of interactive models, such as McClelland and Elman's (1986) TRACE model: Active lexical representations should increase the activation of their component phonemes via top-down excitatory connections. We have begun a second research program that tests another prediction of interactive models: The activation of one word is hypothesized to inhibit the activation of others; this "lexical inhibition" is due to competition between similar lexical items. We are in the early stages of investigating this subtle issue. This work is being done in collaboration with Uli Frauenfelder, now at the University of Geneva (Switzerland); Roger Dowd, a participant in Stony Brook's Research Experience for Undergraduates, assisted with our first experiment in this research program.

According to TRACE the effects of competitive inhibition are due to the combined effects of lexical feedback and lateral inhibition between phoneme units (McClelland and Elman, 1986). This is because TRACE allows the lexical level to excite but not to inhibit the phoneme level whereas it allows phoneme units to inhibit other phoneme units but not to excite them. Therefore, this model predicts indirect inhibition from the lexical level to the phoneme level through the combined effects of positive lexical feedback to the phoneme level and within level inhibition between phoneme units. Interactive models can be contrasted with purely bottom-up, or autonomous, models, such as Cutler and Norris's RACE model (1979). The RACE model does not allow for inhibitory effects. This model states that the two competing procedures (lexical, phonemic) function completely independently so that the lexical code can never affect the computation of the prelexical (phonemic) code.

Our first experiment tested whether lexical inhibition occurs, as hypothesized by TRACE. On each trial, subjects were presented with a visual target (a word or nonword), followed by the dichotic presentation of monosyllabic auditory stimuli. There were four dichotic cases: word-word, word-nonword, nonword-word, and nonword-nonword. In one

condition a word and nonword were presented dichotically. For example, 'gift' might be presented vi...ally, followed by the dichotic presentation of 'gift' and 'kift'; the subjects' primary task was to respond 'same' if they heard the word 'gift', or 'different' if they did not. An additional task was to decide whether they heard a word. In this example, because 'gift' was presented in one ear, 'word' would be the correct response. TRACE predicts that 'gift' will be correctly and rapidly recognized because 'kift' is a nonword and thus is not a very strong competitor. The /k/ will be inhibited by the appropriate phoneme /g/ due to the effects of lexically mediated inhibition. The reverse of this condition is a nonword-word condition in which a nonword ('kift') is the visual target. Here, TRACE predicts that the word will compete for recognition resulting in increased response times. For example, if 'gift-kift' is the auditory stimulus, and 'kift' is the target, 'gift' should increase reaction times due to the effects of competitive inhibition. A third condition involved a nonword-word auditory stimulus. In this condition if 'guisp-kisp' was presented with 'guisp' as the visual target, subjects should respond more quickly than in the 'kift' case, due to the lack of competition (i.e., nonwords have no representation in our mental lexicon, thus lexically mediated inhibition cannot exist; responses will based upon prelexical processes). Finally, a fourth condition involved a word-word auditory stimulus. Here, it was predicted that there will be increased response time latencies. For example, if 'guild-killed' is presented with 'killed' as the visual target, 'guild' will compete with 'killed' for word recognition, increasing response times.

Insert Table 1 Around Here

Table 1 summarizes the stimulus conditions. Critical stimuli were chosen to produce a 2 (target word vs. target nonword) by 2 (nonword competitor vs. word competitor) within subjects design. Out of a total of 80 critical stimulus items, 40 were critical word trials, in which the visually presented target was a word. On 20 of these trials, the subject heard a dichotic word-nonword stimulus, and on 20, a dichotic word-word stimulus. On the other 40 critical trials, the subject received a nonword target. Half of these trials included a word-nonword stimulus, and half involved a nonword-nonword stimulus.

On Filler trials, only a single auditory stimulus was presented. More specifically, the filler word or nonword was presented to the two ears, with a 60 msec delay between the two ears. This allowed a single item to be presented in a way that was phenomenologically similar to the critical dichotic trials. There were 400 Filler items, 284 of which were word items. Those involved 224 (56%) matched word pairs (word target with matching word stimulus) out of a total of 280 (70%) matched pairs, and 60 mismatched pairs. Of the 60 mismatched pairs, 24 had a word target and a rhyming word lure (e.g., target = 'swipe', stimulus - 'ripe'); 6 had a nonword target and a rhyming word lure (e.g., 'clim-rim'); 24 had a word target with a matching onset word lure (e.g., 'wide-wife'); and 6 had a nonword target with a matching onset word lure (e.g., 'shom-shop'). The remaining 116 filler items were nonword items, of which 56 match (nonword target with matching nonword), and 60

mismatch. Of the 60 mismatched pairs, 24 had a word target with a rhyming nonword lure (e.g., 'life-jife'); 6 had a nonword target with a rhyming nonword lure (e.g., 'wibe-zibe'); 24 had a word target with a matching onset nonword lure (e.g., 'rod-rop'); and 6 had a nonword target with a matching onset nonword lure (e.g., 'frid-frip'). The higher proportion of word (versus nonword) fillers was designed to encourage lexical processing.

Critical stimulus pairs began with the stop consonants /b/, /p/, (e.g., 'bob-pob'); /d/, /t/, (e.g., 'dot-tot'); or /g/, /k/, (e.g., 'gift-kift'). Reaction times were compared between stimulus pairs which share the same initial stop consonants and first vowel. For example, a critical comparison might involve 'dent' (in 'dent-tent') versus 'deaf' (in 'deaf-tef'); these share the same initial consonant (/d/) and first vowel (/E/).

Subjects were told that they would be shown a visual cue on a computer monitor and that their task was to respond as quickly as possible without sacrificing accuracy, as to whether what they saw was the 'same' or 'different' from what they heard, and whether or not what they heard was a 'word' or a 'nonword'. On each trial the visual cue was presented for 750 milliseconds, followed by 500 milliseconds of silence before the auditory stimulus was presented. Subjects' first response was made by hitting one or two buttons labeled "same" and "different" on a response console. Their second decision was made by hitting a second pair of buttons labeled "word" and "nonword". The 40 word target, critical item, reaction times of word-word pairs were compared with word-nonword pairs, and the 40 nonword target, critical item, reaction times of nonword-nonword pairs were compared with word-nonword pairs. Two primary measures were analyzed. The first is a comparison of reaction times for 'same' responses for word target items, and 'same' responses for nonword target items. Therefore, reaction times of 'same' responses to word-word pairs were compared with 'same' responses of nonword-word pairs for word target items. Similarly, reaction times of 'same' responses to nonword-nonword pairs were compared with 'same' responses of nonword-word pairs for nonword target items. The second analysis is a comparison of reaction times for the lexical decision of "word or nonword", in which comparison sets were the same as above. A two-way analysis of variance (target type x competitor type) was run on reaction times for both decisions.

Insert Figure 10 Around Here

Figure 10 shows the mean reaction times for word and nonword targets, with word and nonword competitors, on the primary (matching) judgment. The main effect for type of target (word or nonword) was significant, with the nonword target mean (1173 msec) being significantly higher than the word target mean (974 msec). Contrary to our expectation, no main effect was found for competitor type. However, there was a significant interaction between target and competitor. As predicted by TRACE, reaction times for word target items with a word competitor had a higher mean (1031 msec) than nonword competitor items (916 msec). However, for nonword target items with a word competitor there was a lower

mean (1173 msec) than for nonword competitor items (1199), which was contrary to TRACE's prediction. These results suggest that TRACE's assumption of lexical competition is correct, but that the assumption of top-down mediated phonemic inhibition is questionable.

Insert Figure 11 Around Here

Figure 11 presents the reaction times in the lexical decision task. The main effect of competitor type was significant, with the word competitor mean (541 msec) being greater than the nonword competitor mean (444 msec). There was no main effect for target type or interaction of target type and competitor type. Reaction times for word target items with a word competitor had a higher mean (505 msec), than nonword competitor items (415 msec). For nonword target items, word competitors yielded a higher mean (576 msec) than did nonword competitors (472 msec), but the difference was not significant.

Overall, the results found were congruent with TRACE's prediction of competitive inhibition. As hypothesized, reaction times were longer in the word competitor conditions. For 'same' responses when there was a word target with a word competitor, reaction times were significantly higher than when there were nonword competitors. For lexical decision responses, with both word and nonword targets, reaction times were significantly higher when there was a word competitor than with a nonword competitor.

Although our first experiment's results look promising, much work remains to be done. There are two clear priorities, one theoretical and one methodological. Theoretically, the vastly different results (for the primary matching judgment) for word versus nonword targets need to be understood. If this difference is real, it will provide very important constraints in specifying how lexical activation occurs. Further experiments will be needed to determine the conditions under which lexical inhibition occurs.

Methodologically, the challenge is to develop tests which provide a more natural approximation to normal listening. In our first experiment, we used dichotic presentation to assure that competition could occur: We actually presented a potential competitor. We plan to construct monaural (or binaural) tests in which a single item can generate lexical competition. For example, we can synthesize a token with a VOT value that is intermediate between /g/ and /k/ in a token that could be heard as "gift" or "kift"; the same onset could be used in the other pairs of that stimulus set (e.g., the word-word case of "guild-killed", or the nonword-nonword "gisp-kisp"). These stimuli will provide a better test of whether lexical competition occurs when a single token is heard. This line of research has the potential to provide important insights into how the representations of a speech signal interact.

<u>C. A Test for Lexical Activation</u>. In the two preceding sections, we have described experiments that test differing predictions of interactive models (e.g., TRACE) and

autonomous models (e.g., RACE). In collaboration with Mark Pitt (at Ohio State), we have been testing another prediction that differentiates the two model classes. These experiments have used the phoneme monitoring task, in which subjects rapidly indicate when a pre-specified target phoneme occurs. Prior research has shown that under some conditions, subjects detect phonemes more rapidly in words than in pseudowords. At a general level, both model classes can account for this lexical advantage. In TRACE, the enhanced detection in words is due to the top-down activation of a word's component phonemes. In RACE, any lexical advantage is attributed to the fact that both lexical codes and prelexical (phonemic) representations get computed for words, with responses generated from whichever code is available first (hence the name, RACE). The distribution of response times for words is therefore due to a mix of the two processes, while for pseudowords only the prelexical codes are usable (since pseudowords by definition do not have lexical representations). Thus for RACE, any lexical effect (faster word RT than pseudoword RT) must be due to the availability of post-lexical access codes.

The basis for our recent experiments is a study by Frauenfelder, Segui, and Dijkstra (1990). Frauenfelder et al. capitalized on the fact that there is now a good body of evidence to support the notion that lexical access occurs when enough of a word has been presented to distinguish it from all other possible words. The point at which this occurs, called the "uniqueness point" (UP), has been shown to be closely related to listeners' recognition of words (e.g., Marslen-Wilson, 1984). It follows from this that if autonomous models like RACE are correct, then lexical effects should only appear for target phonemes after the UP. Before this point, the listener cannot extract the phonological information needed to identify the target phoneme, because the correct lexical entry has not yet been identified. As Frauenfelder et al. put it, they sought to determine whether lexical effects "emerge before or after the uniqueness point as TRACE and RACE models, respectively, predict" (p. 80).

To test the theories, they had subjects monitor for phoneme targets in words and in very closely matched pseudowords, probing both before and after the uniqueness point of the words. In contrast to RACE, interactive models predict that targets in words should be detected faster than targets in pseudowords, even before the UP. This is because in models like TRACE, there is a continuous increase in the activation level of words that are consistent with the input signal, even before the UP. Moreover, there is continuous interaction between lexical and sublexical levels. Therefore, such models predict that a lexical effect should be observable even before the UP.

Unfortunately, Frauenfelder et al.'s results were not decisive: While there was no lexical advantage in token-initial position, and there was such an effect after the UP, the critical probes beyond initial position but before the uniqueness point produced a statistically marginal 17 msec word advantage.

Our recent experiments have been designed to resolve this marginal situation. We believe that two aspects of Frauenfelder et al.'s experiment might have reduced the strength of any underlying lexical effect. One problem is that they may have matched their words

and pseudowords too well: By using pseudowords that were very similar to the original words, they may have actually engaged lexical representations in the perception of the pseudowords. In models like TRACE, pseudowords are processed by partial matches to existing lexical representations. If the very tightly matched pseudowords did activate lexical representations, then it would be very difficult to find any difference between them and true lexical items that activated the same sort of representations.

A second factor that could affect the observed results is how much time was available for activation of lexical codes to occur. In TRACE, such activation takes time, and if the targets came relatively early in a word, there might not have been sufficient activation time for the lexical feedback to occur. These considerations led us to design a stimulus set in which (1) the pseudowords, though matched to the words, nevertheless were sufficiently nonlexical to make lexical activation unlikely, and (2) the location of the pre-UP target was manipulated to see whether providing more activation time would lead to a reliable lexical effect for targets before the UP.

Insert Figure 12 Around Here

Figure 12 illustrates the two critical features of our stimuli. The top of the figure shows the word "circumvent" and its matched pseudoword "silpumvent". The four possible probed positions in each are underlined. Note that even though targets in the two stimuli are either identical or nearly so (i.e., members of the some phone class), the pseudoword is not likely to engage a lexical representation. In fact, each pseudoword was chosen so that by the time the first phoneme of the second syllable was heard, the stimulus would mismatch all words in the English language (in this example, there are no English words that begin with /sIlp/.

The bottom of Figure 12 illustrates how we manipulated the activation time available for a word. Half of our words were chosen to have early uniqueness points: on average, only 9 words existed that matched the first syllable of an "Early Unique" word. In contrast, our "Late Unique" words were chosen to have many "neighbors": There were almost 300 English words that shared the first syllable of each Late Unique word. Therefore, we could choose targets like the /m/ in "circumvent" that were still before the UP, but for which there would be a substantial amount of activation time. In contrast, the /c/ in "sarcastic" comes very early in the word, even though it is just before the UP, just as the /m/ in "circumvent" is.

We predicted that if interactive models are correct, then our Late stimuli should produce a lexical advantage: faster detection of targets than in matched pseudowords. This would occur because (1) the pseudowords were chosen to prevent lexical activation, and (2) the targets in Late words occur late enough to allow activation effects to emerge. In contrast, interactive models predict no lexical advantage for the Early stimuli, because these

would not provide sufficient activation time. Autonomous models predict no pre-UP lexical advantage in either case. Both model types predict no lexical effect for token-initial targets, and both predict an effect for targets after the UP.

In our first experiment, we had subjects detect targets in words and pseudowords, with the targets occurring either token-initially, or in the BUP (before uniqueness point) positions. There were 36 Early Unique words, and 36 Late, each with a matched pseudoword. On each trial, subjects saw a target on a CRT (e.g., "R"), and then heard a word or a pseudoword. They were told to respond by pushing one of two buttons, to indicate target presence or absence. Speed of response was stressed.

Insert Figures 13 and 14 Around Here

Figure 13 shows the average reaction times for phoneme detection, for targets in the Early Unique words and in their matched pseudowords. Error rates are shown in parentheses, for each condition. The results replicate those of Frauenfelder et al.: no significant differences were found between RT to targets in words and pseudowords, in either probed position. These results are consistent with both autonomous and interactive models, assuming the latter require more activation time than the Early words provide.

The results for the Late Unique words were quite different, as shown in Figure 14. In this case, there was a reliable RT difference for probes before the uniqueness point. In particular, the effect for BUP2 is as predicted by the interactive models, and is counter to the predictions of autonomous ones. The significant effect at BUP1 was unexpected, but appears to be real -- it has shown up in four different experiments in this line of research. We will discuss its implications shortly.

Before doing so, we will consider a second experiment in this line that tests whether activation time is an important factor in finding lexical effects. In our second experiment, we used a speech expansion algorithm to slow down the Early Unique words and their matched pseudowords. We stretched these tokens so that the BUP1 targets now occurred at the same point in time as the BUP2 targets in the (unstretched) Late Unique words. If the lexical effect found in our first experiment is due to lexical activation, then the previously ineffective Early words should now have enough time for activation to build sufficiently to produce a lexical advantage.

Insert Figures 15 and 16 Around Here

Figure 15 shows the results for the Early Unique words and their matched pseudowords. As predicted by interactive models, pre-uniqueness point targets in words now

were detected more quickly than the corresponding targets in pseudowords. Thus, providing additional activation time by slowing down the words was sufficient to produce a lexical advantage. Figure 16 shows the results for the Late Unique words in this experiment. This condition was identical to the Late condition of Experiment 1, and produced similar results: a significant word advantage of BUP1, and an advantage of BUP2 (in this case, not significant).

We have run a third experiment in this series in which we compressed the Late Unique words and their matched pseudowords. Note that this deprives the words of activation time, but preserves the amount of bottom-up support for a lexical item. For example, the /m/ in "circumvent" is still preceded by five phonemes that are all consistent with the lexical item. The compression experiment thus allows us to test whether activation time is necessary, or whether sufficient bottom-up support can take its place.

Insert Figures 17 and 18 Around Here

As Figure 17 shows, time is apparently not necessary -- enough bottom-up support will suffice to produce a lexical advantage. For both BUP probed positions in the Late words, a reliable lexical advantage was found. In fact, for reasons that are not yet clear, there was even a lexical advantage in word-initial position. The results from the (uncompressed) Early words and pseudowords, shown in Figure 18, include a similar effect in initial position, but more importantly, include a replication of the lack of effect at BUP1 (as in Experiment 1).

The results for the three experiments are thus generally consistent with the predictions of TRACE: A lexical advantage will occur when there is either sufficient activation time or sufficient bottom-up support for a word, even before the uniqueness point. This is precisely the effect that Frauenfelder et al. argued was a critical point for differentiating interactive models from autonomous ones.

There is one recurring result that needs further comment, and further research. As is clear in several of the figures, we have consistently found a reliable RT advantage at BUP1 for the Late words relative to their pseudowords. This effect is not predicted by either model class, yet it is very robust. At this point, we believe that this effect reflects sublexical processes. In particular, we suggest that very high frequency syllables may have their own representations, and these representations can support phoneme detections in much the same way that lexical ones can. Note that the first syllables of Late Unique words were all very high frequency -- on average, almost 300 words shared the first syllable.

At this point, this suggestion of a sublexical basis for the Late BUP1 effect is quite speculative. However, we have been able to rule out one important alternative account so far. This alternative is the possibility that the target phonemes in BUP1 position just

happened to be more salient than the corresponding targets in the matched pseudowords. To test this artifactual possibility, we removed the initial phoneme from each word and pseudoword, and re-ran our experiment. Note that such a "beheading" effectively eliminates the high frequency syllable (e.g., "sub" in "substantial" is now "ub", a low-frequency syllable), but would preserve any artifactual differences in the target phonemes later in the syllable.

Insert Figures 19 and 20 Around Here

Figure 19 presents the RT data for these beheaded stimuli made from Late Unique words and their matched pseudowords. There is now no hint of a lexical advantage in either BUP position. If the previous lexical effects were due to artifactual stimulus differences, the RT difference should have remained. That it did not suggests that our sublexical account is plausible. For the stimuli made from Early Unique words, Figure 20 shows no lexical effect. This, of course, is not surprising, as there was no such effect for the full stimuli.

Taken together, the experiments in this line of research provide support for interactive models like TRACE. They do so by finding a lexical advantage before the uniqueness point, a result that is difficult to accommodate in autonomous models. The experiments indicate that either adequate activation time, or sufficient bottom-up support, can produce a lexical advantage. In addition, we have found an unexpected advantage that does not appear to be lexically based. Instead, we suggest that the result reflects activation of high frequency sublexical representations. More research is needed to firm up the lexical findings, and to begin to understand the hypothesized sublexical effects.

IV. List of Publications

Pitt, M. A., and Samuel, A. G. An empirical and meta-analytic evaluation of the phoneme identification task. Journal of Experimental Psychology: Human Perception and Performance, in press.

V. Personnel

Principal Investigator: Arthur G. Samuel, Associate Professor of Psychology at the State University of New York at Stony Brook. Ph.D. from University of California, San Diego, 1979.

Senior Research Specialist: Donna Kat, B.A. in Psychology from University of California, San Diego, 1979.

Graduate Student: Lee Wurm. Mr. Wurm joined our lab during the summer of 1992, and has been studying many of the issues reviewed in Section III of this report.

IV. List of Interactions (coupling activities)

Presentations

Samuel, A. G. Perceptual restoration, perceptual bias, priming, and pseudowords: Insights from a newer methodology. Presented at the Psychonomic Society, San Francisco, November 1991.

Pitt, M. A., and Samuel, A. G. Is auditory word recognition serial or interactive? Presented at the Psychonomic Society, San Francisco, November 1991.

Samuel, A. G. Probing words and pseudowords: Evidence for interactive activation models of word recognition. Presented at the Max Planck Institute for Psycholinguistics, Nijmegen (Netherlands), August 1991.

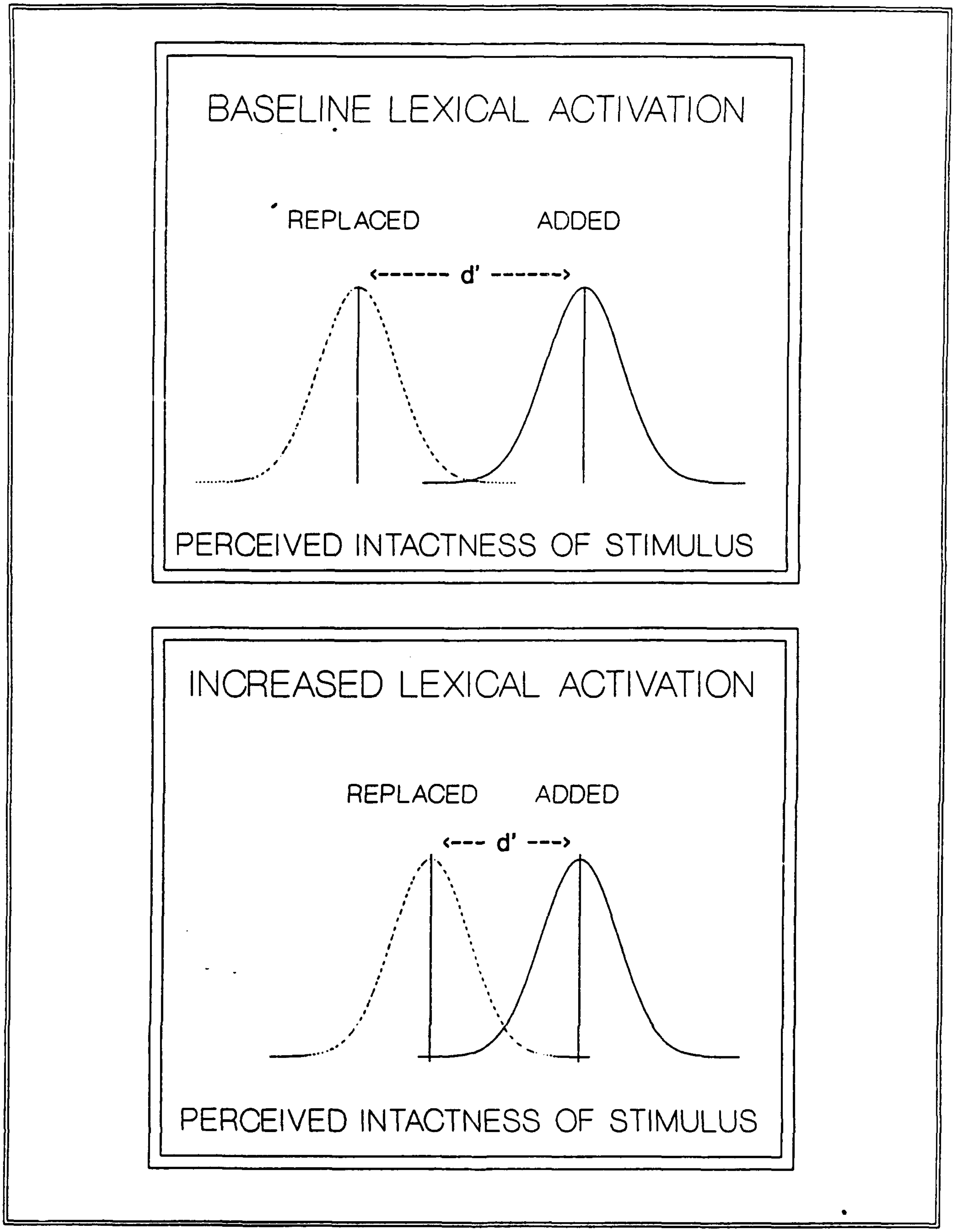
Other Interactions:

The P.I. was invited to be a Visiting Scholar at the Max Planck Institute (Nijmegen, The Netherlands) in August of 1992. This allowed for many useful interactions with European scientists. In addition, the P.I. has been on the Editorial Boards for Cognition, Memory and Cognition, and the Journal of Experimental Psychology: Human Perception and Performance. In this capacity, he has reviewed many papers. Additional reviews were done for NSF and for several other journals. The P.I. also has joined the Perception and Cognition Review Panel for NIMH, which is a very rich source for interaction with top scientists from around the country.

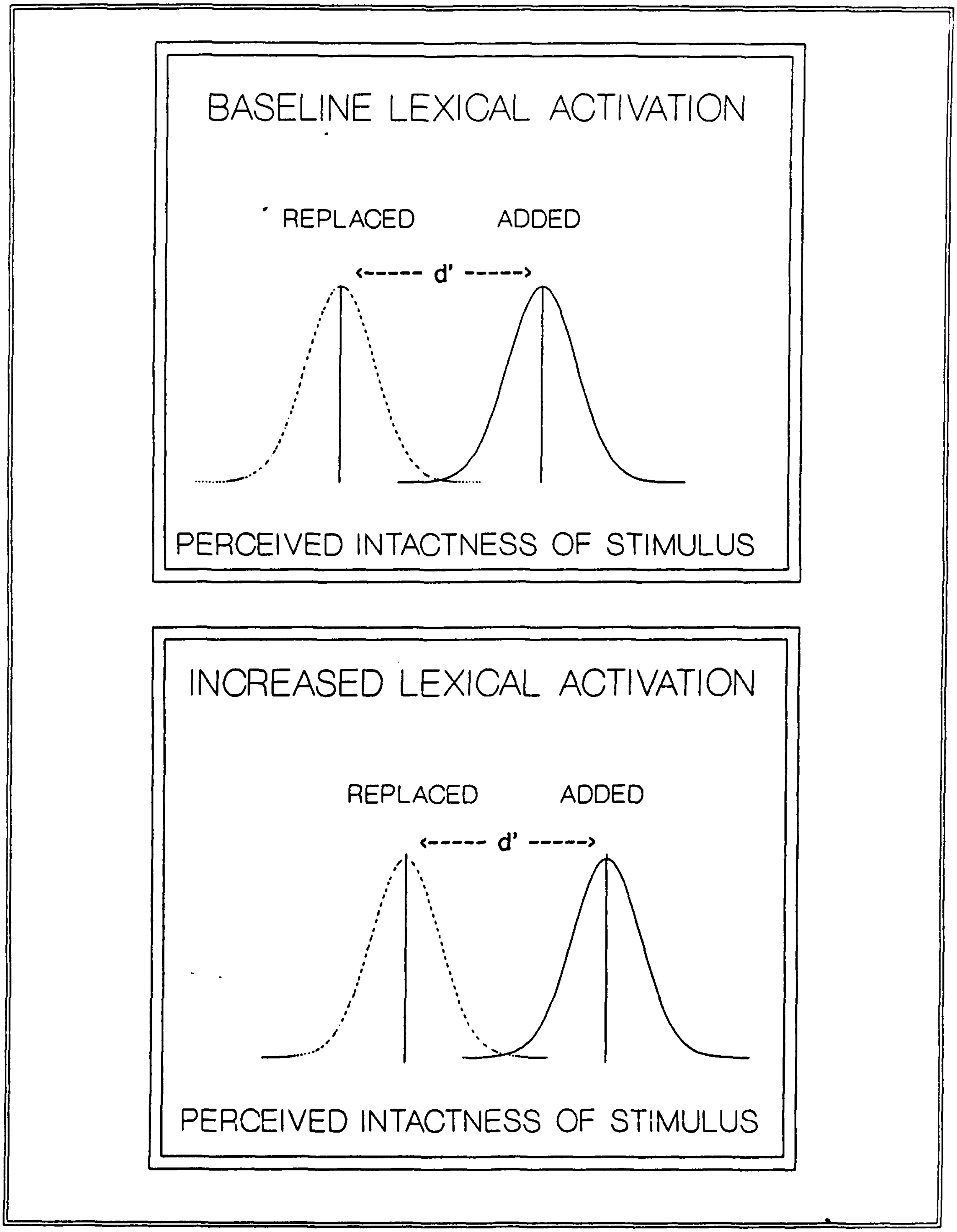
Table 1
Examples of Critical and Filler items

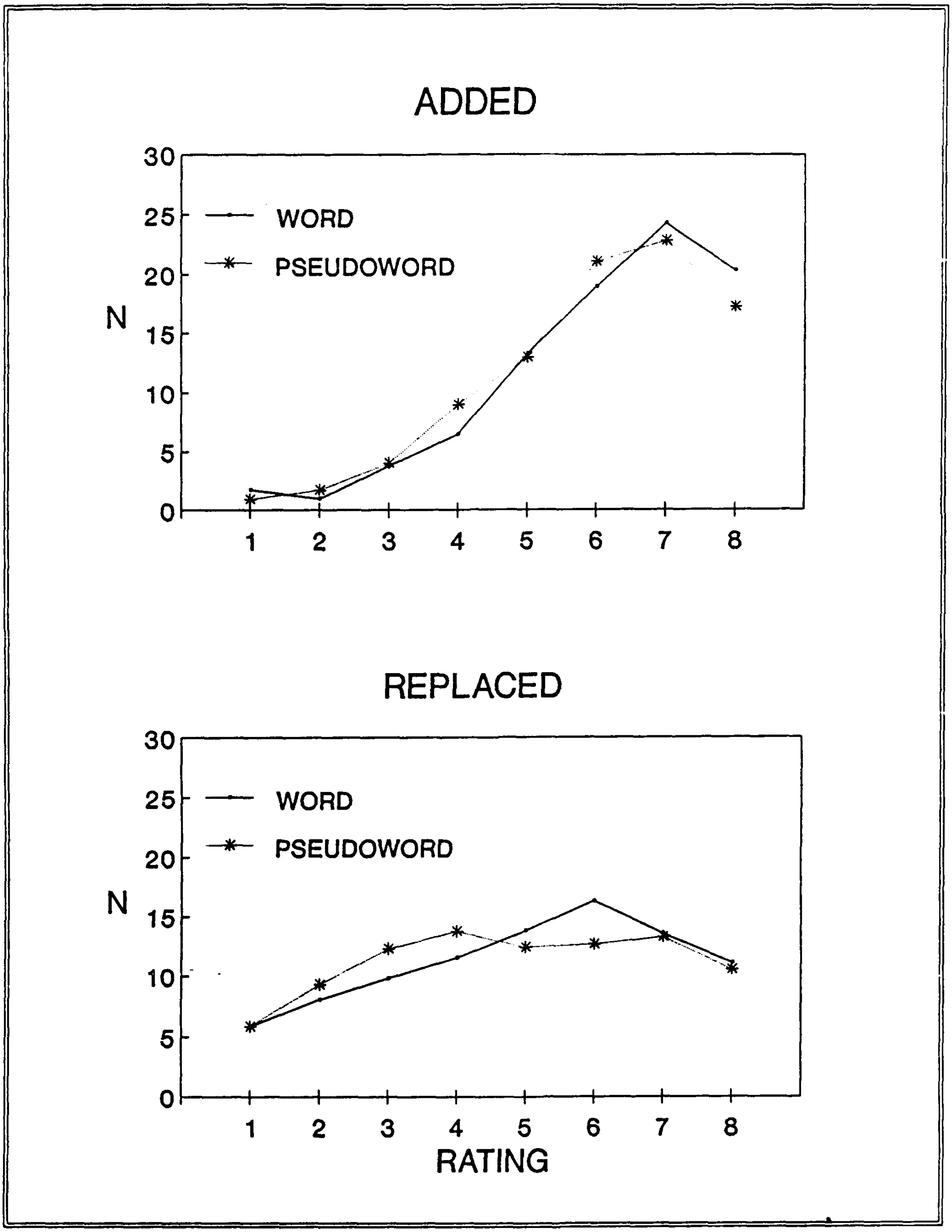
# of stimuli	Visual target	example	auditory stimulus	example
Critical				
20	W	dent	W-W	dent-tent
20	W	deaf	W-NW	Deaf-tef
20	NW	tib	NW-NW	dev-tev
20	NW	tig	W-NW	dend-tend
Filler				
24	W	make	R-W lure	fake
24	W	chart	R-NW lure	zart
6	NW	wibe	R-NW lure	zibe
6	NW	soach	R-W lure	roach
224	W	ride	M-W	ride
56	NW	lisk	M-NW	lisk
24	W	groan	MO-W lure	gross
6	NW	clid	MO-W lure	cliff
24	W	name	MO-NW lure	naist
6	NW	murch	MO-NW lure	murn

W = word, NW = nonword, R - rhyming, M = matching, MO = matching onset

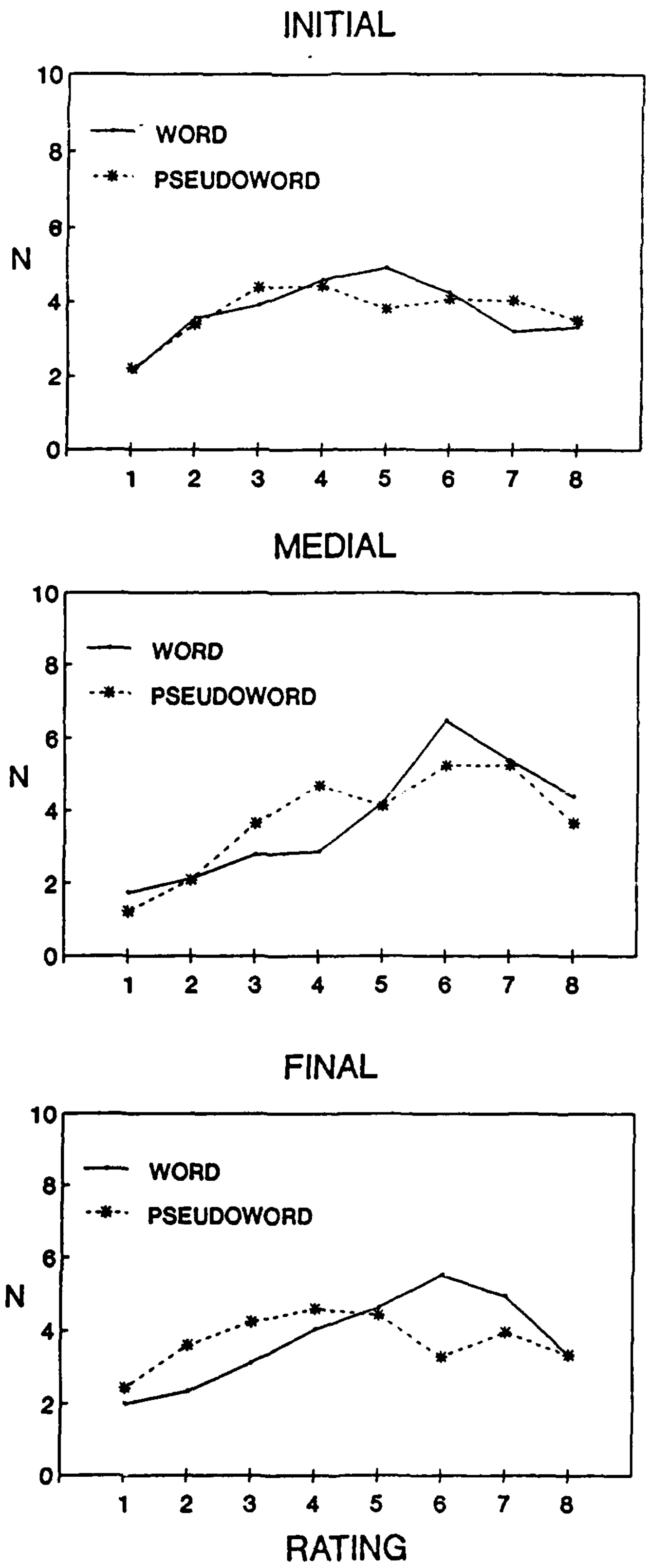

BASELINE LEXICAL ACTIVATION
REPLACED
ADDED
d'
PERCEIVED INTACTNESS OF STIMULUS
INCREASED LEXICAL ACTIVATION
REPLACED
ADDED
d'
PERCEIVED INTACTNESS OF STIMULUS

2

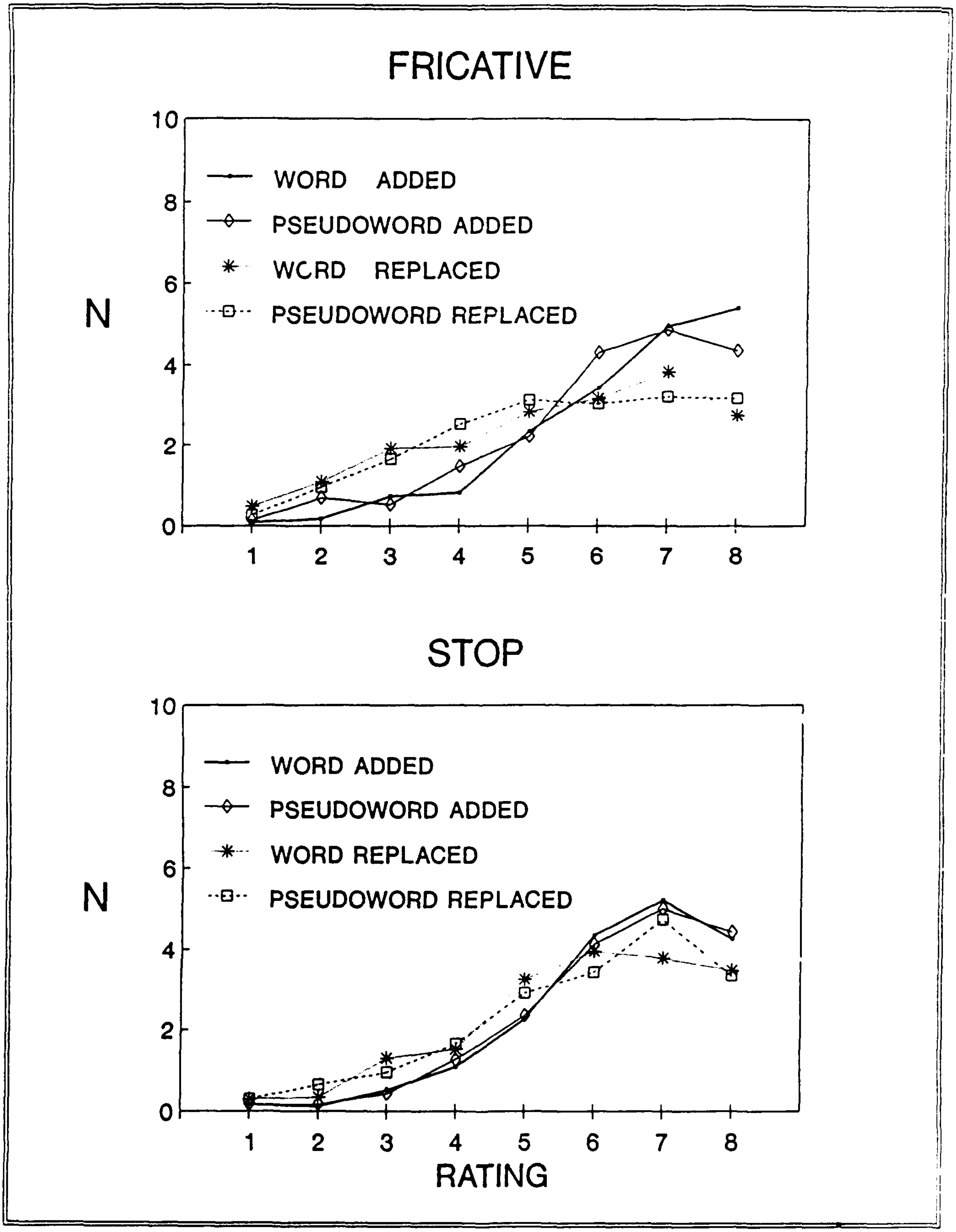

BASELINE LEXICAL ACTIVATION
REPLACED
ADDED
d'
PERCEIVED INTACTNESS OF STIMULUS
INCREASED LEXICAL ACTIVATION
REPLACED
ADDED
d'
PERCEIVED INTACTNESS OF STIMULUS

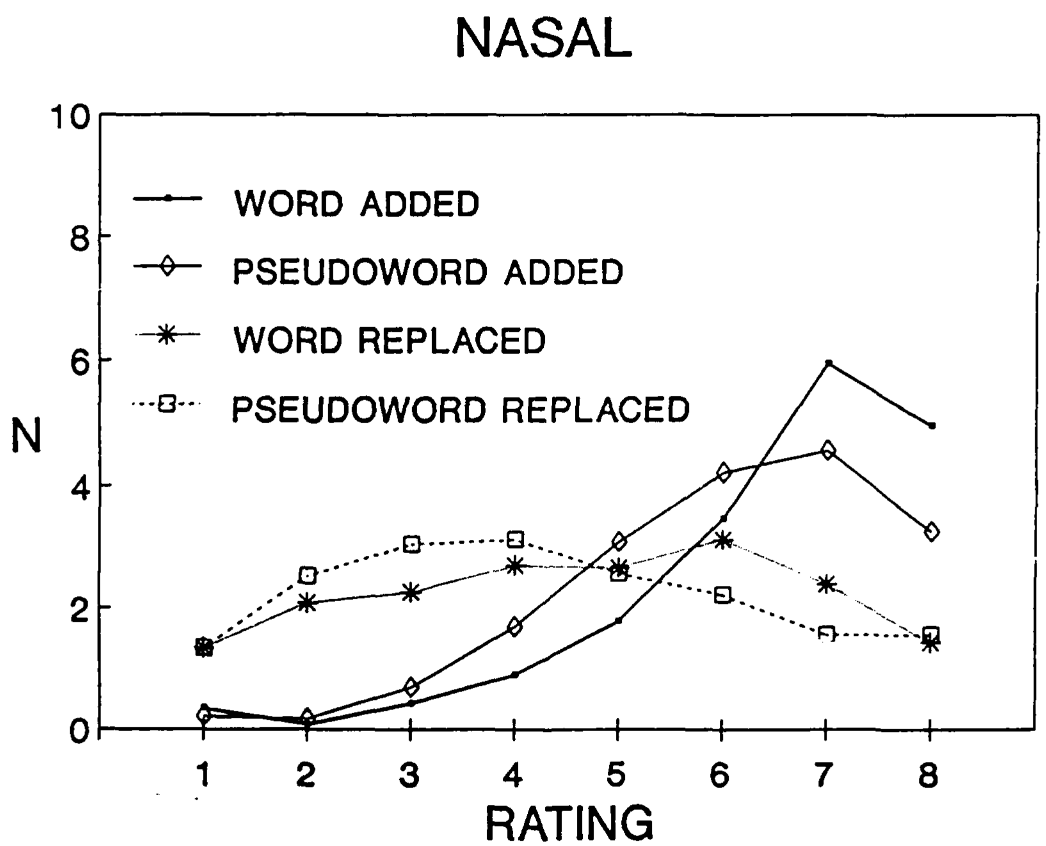


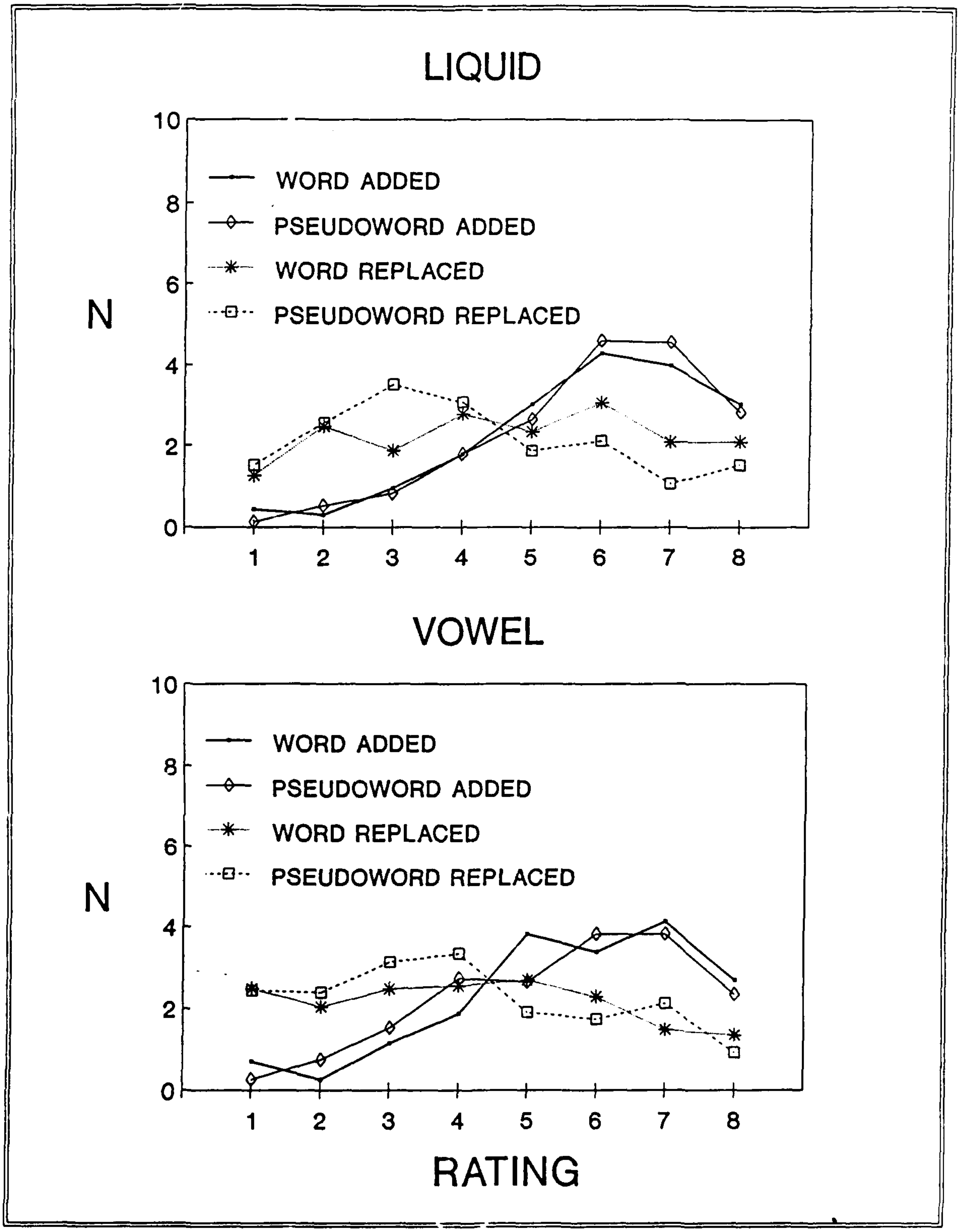
ADDED
30
25
20
15
10
5
0
N
WORD
PSEUDOWORD
1
2
3
4
5
6
7
8
REPLACED
30
25
20
15
10
5
0
N
WORD
PSEUDOWORD
1
2
3
4
5
6
7
8
RATING



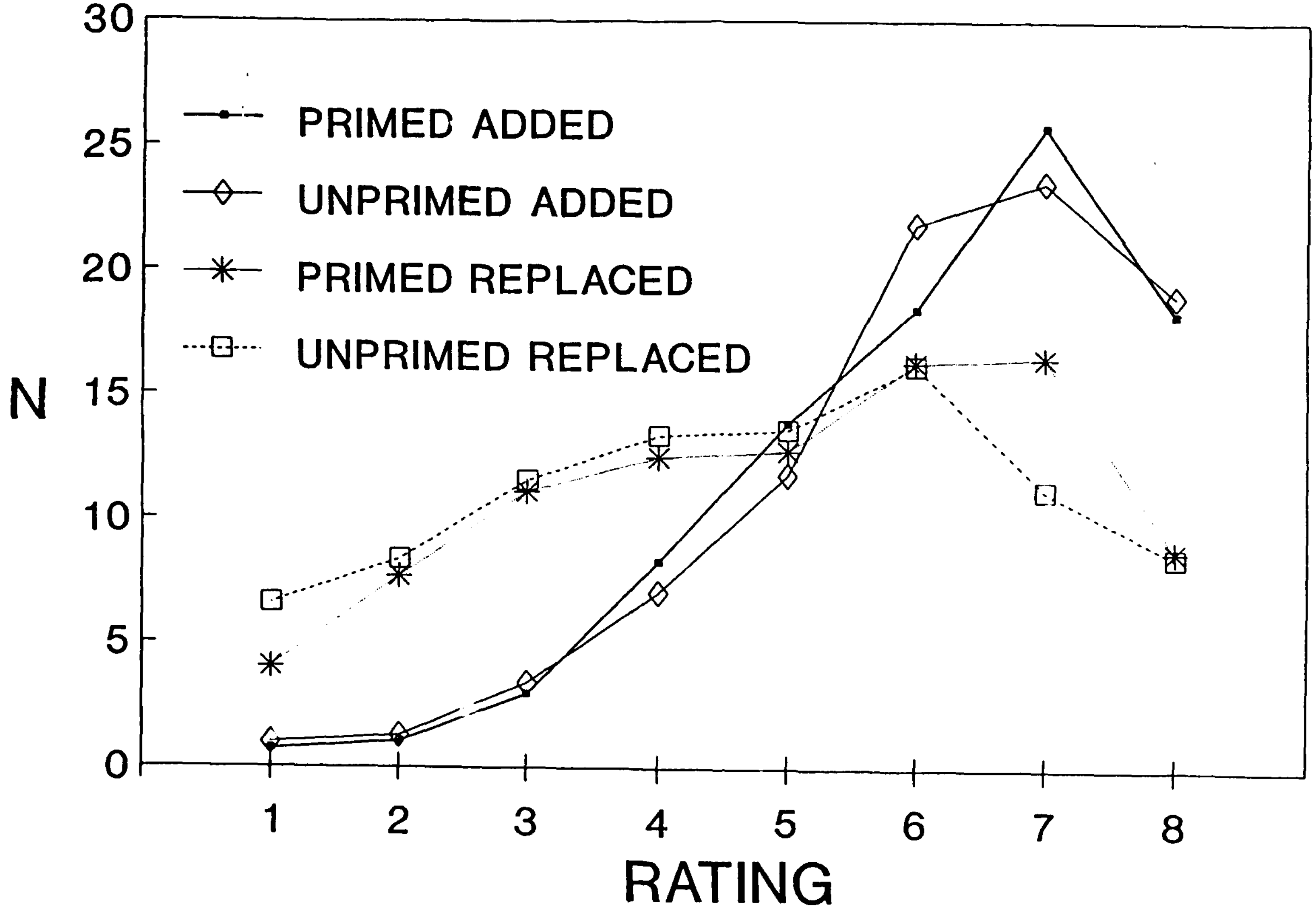
INITIAL
WORD
PSEUDOWORD
N
0
2
4
6
8
10
1 2 3 4 5 6 7 8
MEDIAL
WORD
PSEUDOWORD
N
0
2
4
6
8
10
1 2 3 4 5 6 7 8
FINAL
WORD
PSEUDOWORD
N
0
2
4
6
8
10
1 2 3 4 5 6 7 8
RATING


FRICATIVE
WORD ADDED
PSEUDOWORD ADDED
WORD REPLACED
PSEUDOWORD REPLACED
N
0
2
4
6
8
10
1
2
3
4
5
6
7
8
STOP
WORD ADDED
PSEUDOWORD ADDED
WORD REPLACED
PSEUDOWORD REPLACED
N
0
2
4
6
8
10
1
2
3
4
5
6
7
8
RATING

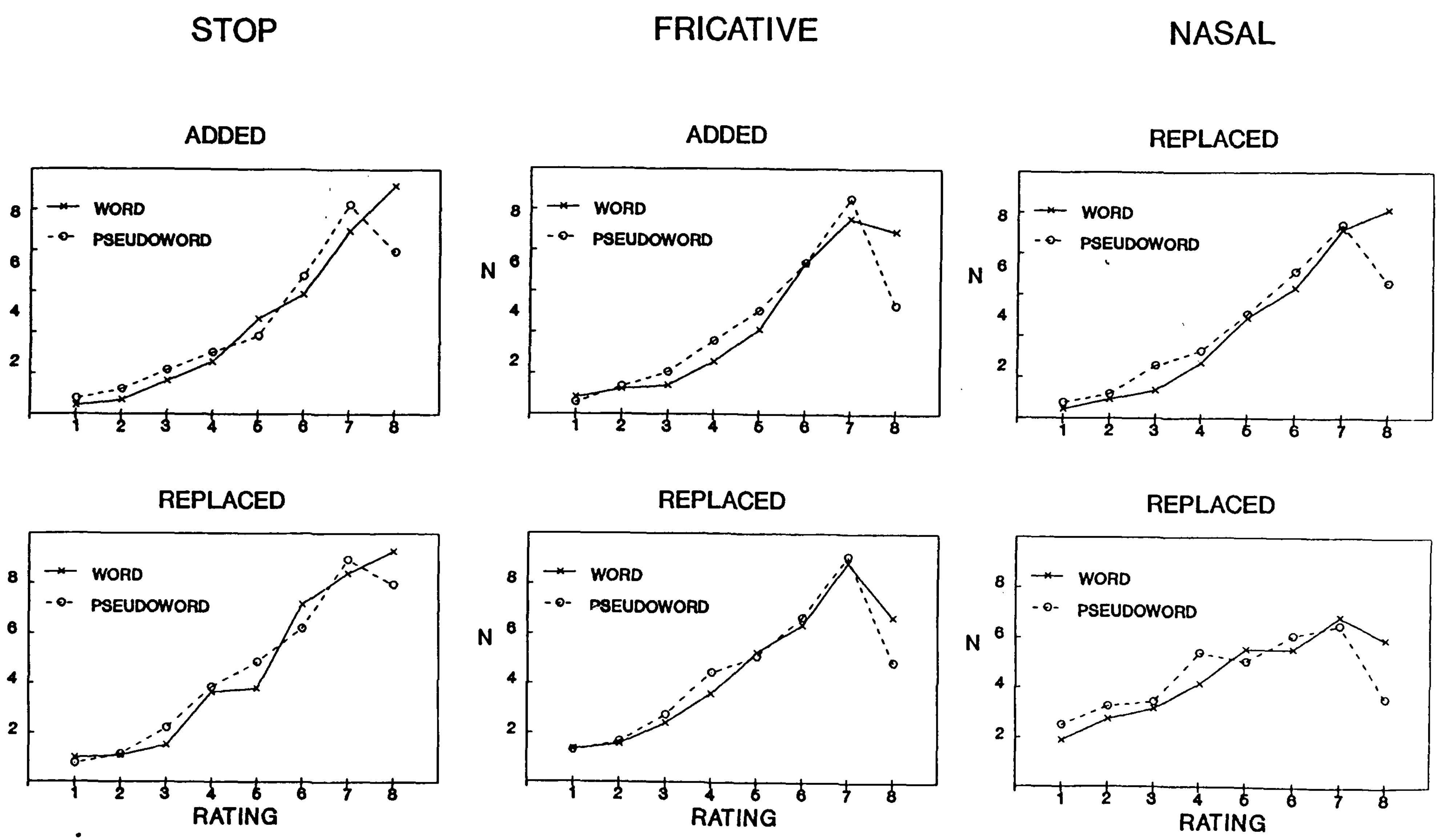

NASAL
10
8
6
4
2
0
N
WORD ADDED
PSEUDOWORD ADDED
WORD REPLACED
PSEUDOWORD REPLACED
1
2
3
4
5
6
7
8
RATING


LIQUID
10
8
6
4
2
0
N
WORD ADDED
PSEUDOWORD ADDED
WORD REPLACED
PSEUDOWORD REPLACED
1
2
3
4
5
6
7
8
VOWEL
10
8
6
4
2
0
N
WORD ADDED
PSEUDOWORD ADDED
WORD REPLACED
PSEUDOWORD REPLACED
1
2
3
4
5
6
7
8
RATING

PRIMING

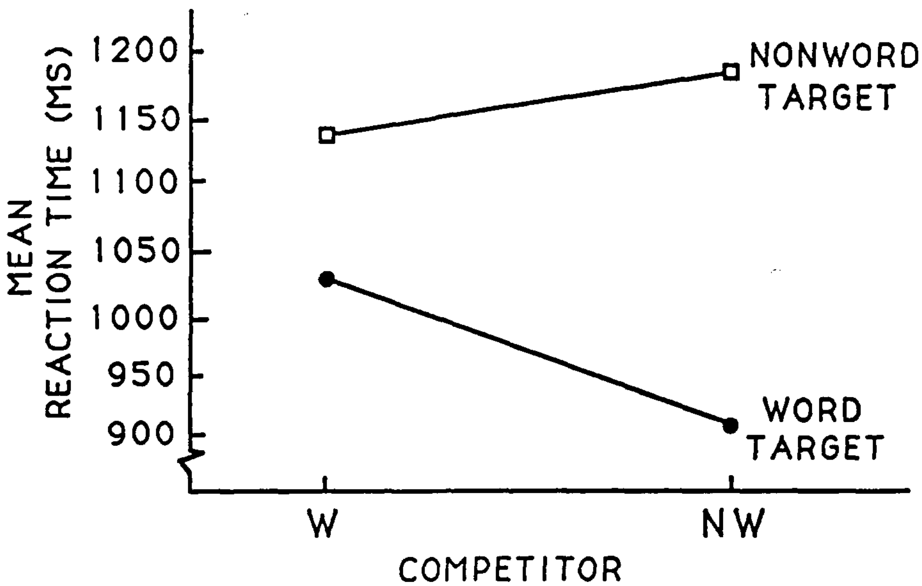

PRIMED ADDED
UNPRIMED ADDED
PRIMED REPLACED
UNPRIMED REPLACED
N
0
5
10
15
20
25
30
1
2
3
4
5
6
7
8
RATING

8

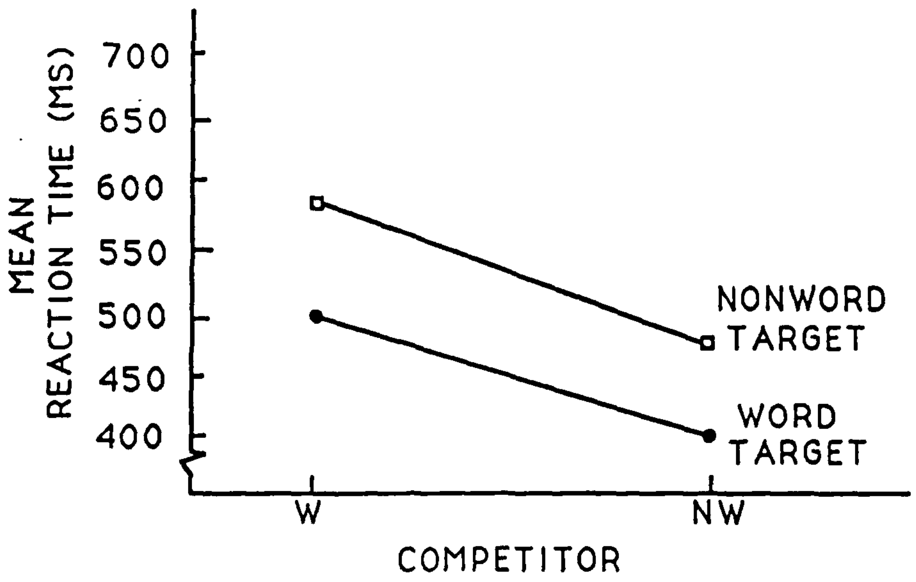


STOP
FRICATIVE
NASAL
ADDED
ADDED
REPLACED
REPLACED
REPLACED
REPLACED
WORD
PSEUDOWORD
N
RATING

"SAME" RESPONSES



LEXICAL DECISION RESPONSES



Condition:

Nonwords S I L P U M V E N T

Words C I R C U M V E N T

WI BUP1 BUP2 AUP

(Before Uniqueness Point)

(Word Initial) (After Uniqueness Point)

(uniqueness point)

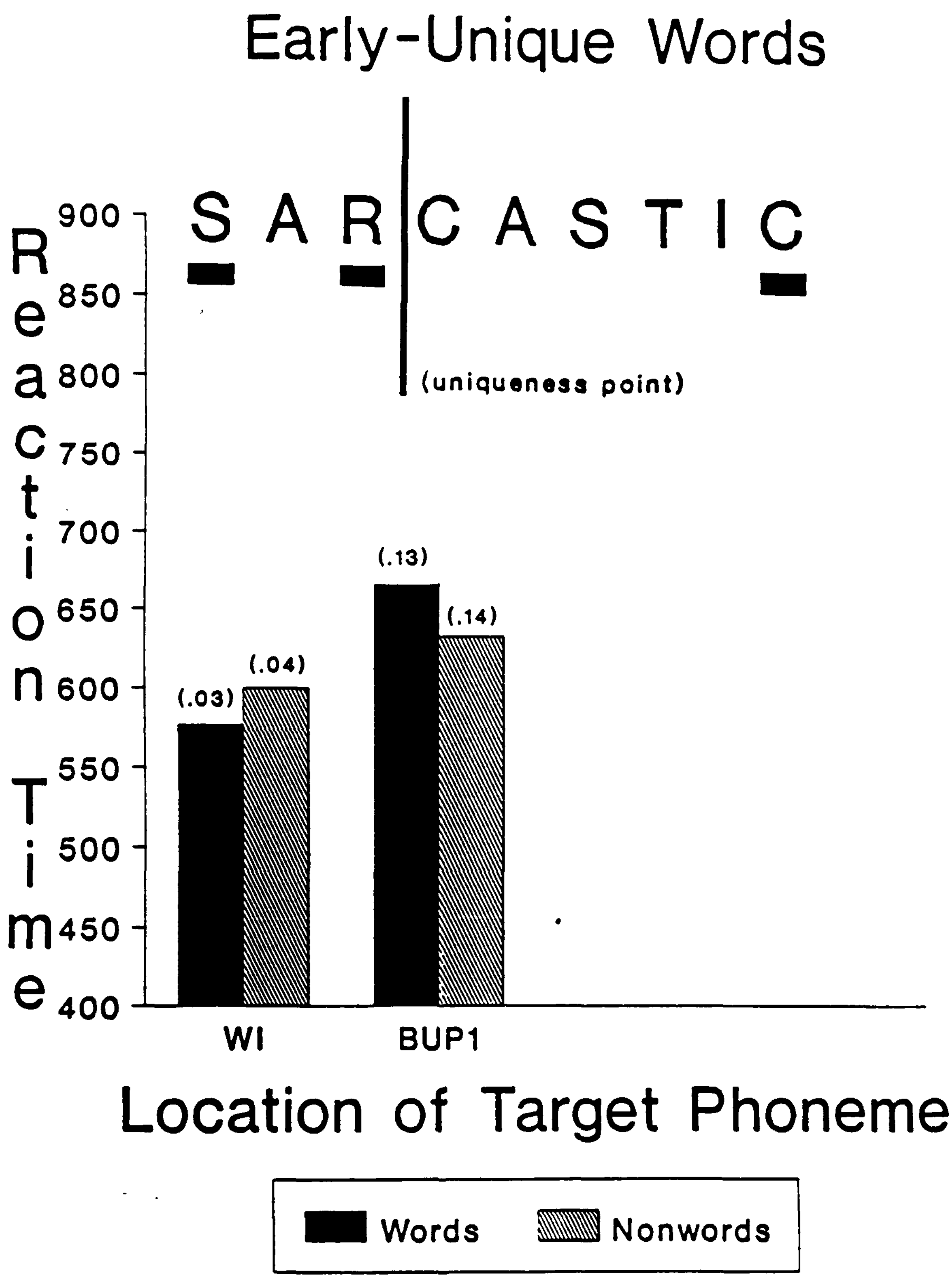
Condition:

Early-Unique S A R C A S T I C

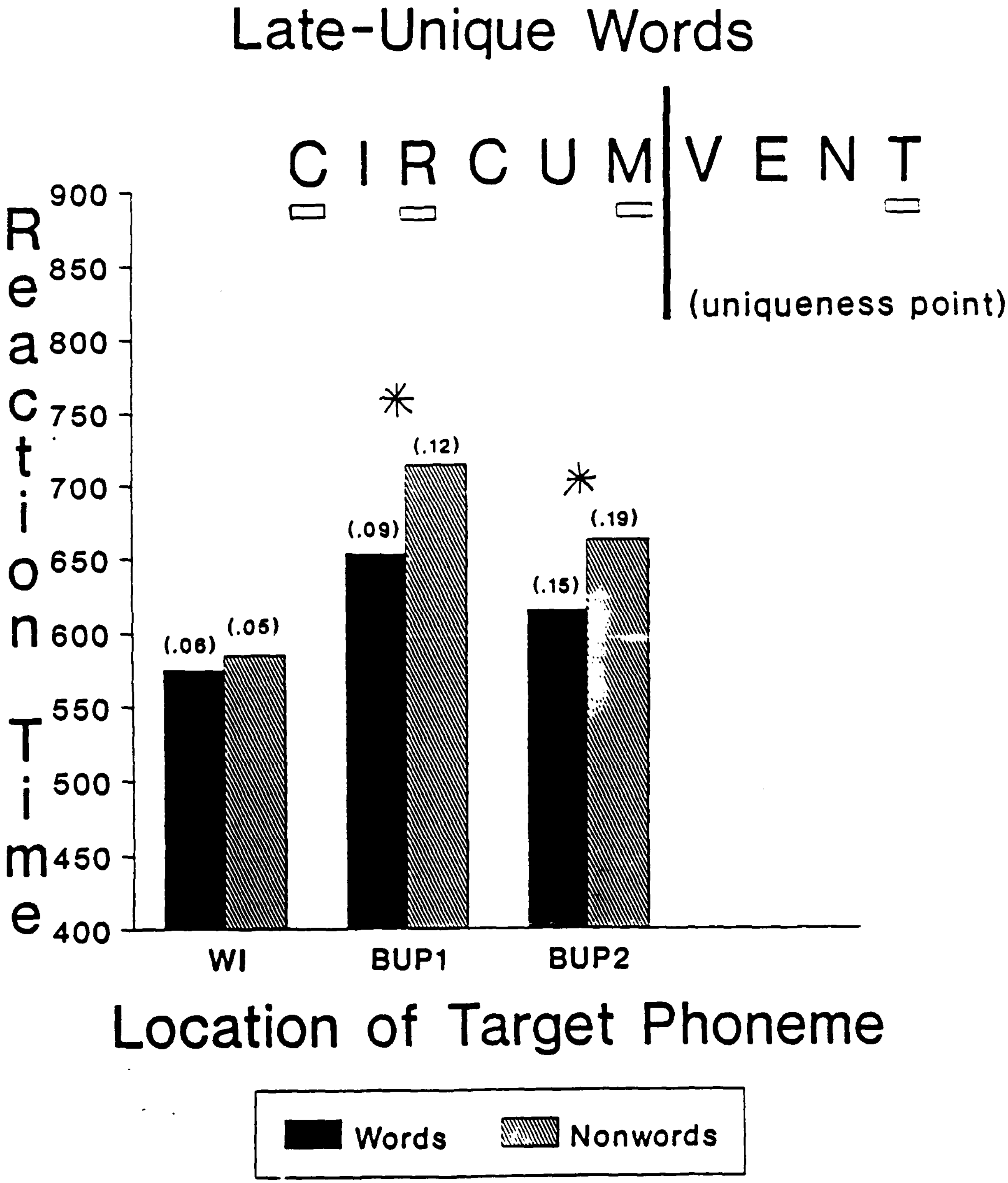
Late-Unique C I R C U M V E N T

WI BUP1 BUP2 AUP

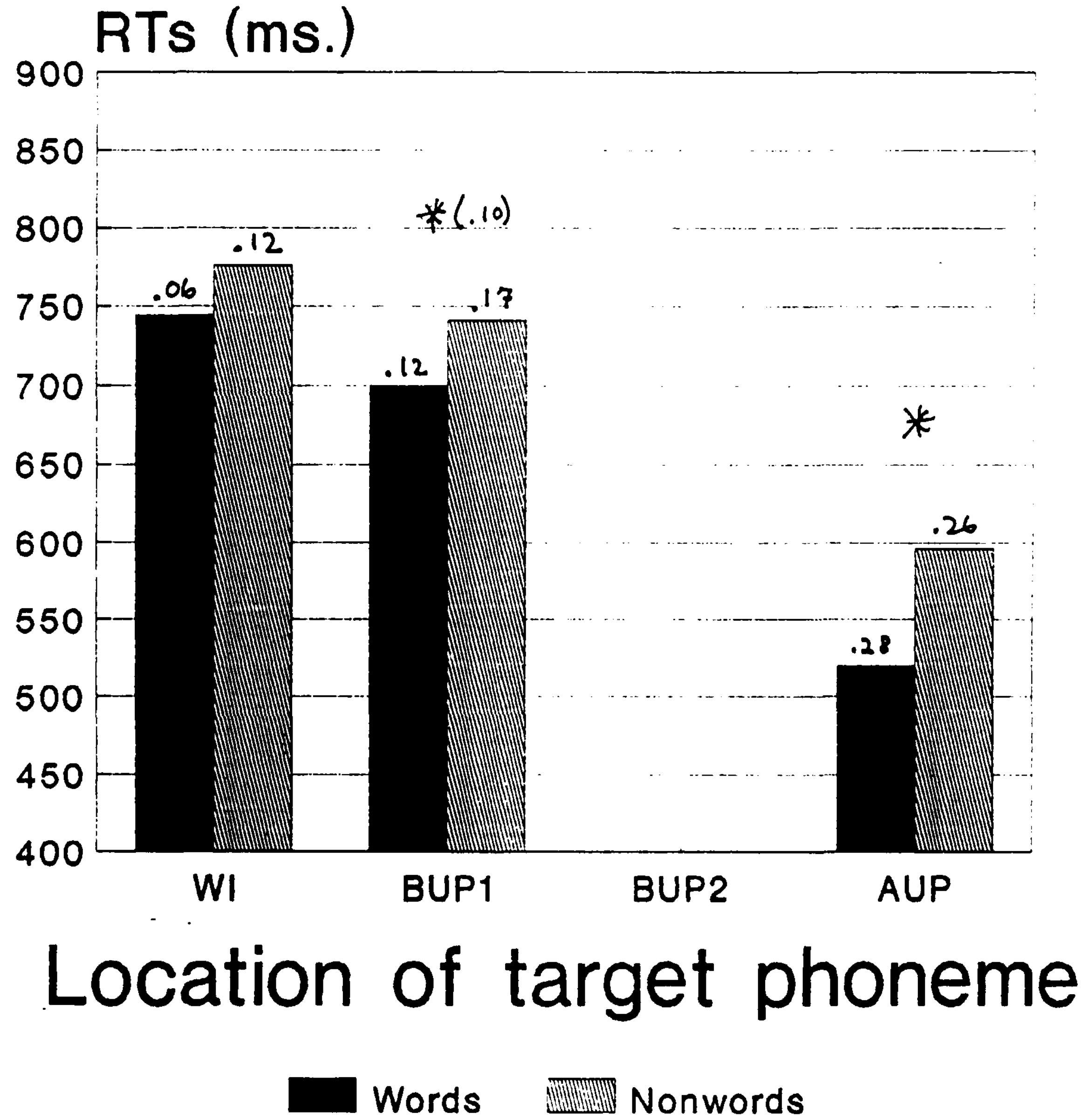
(uniqueness point)



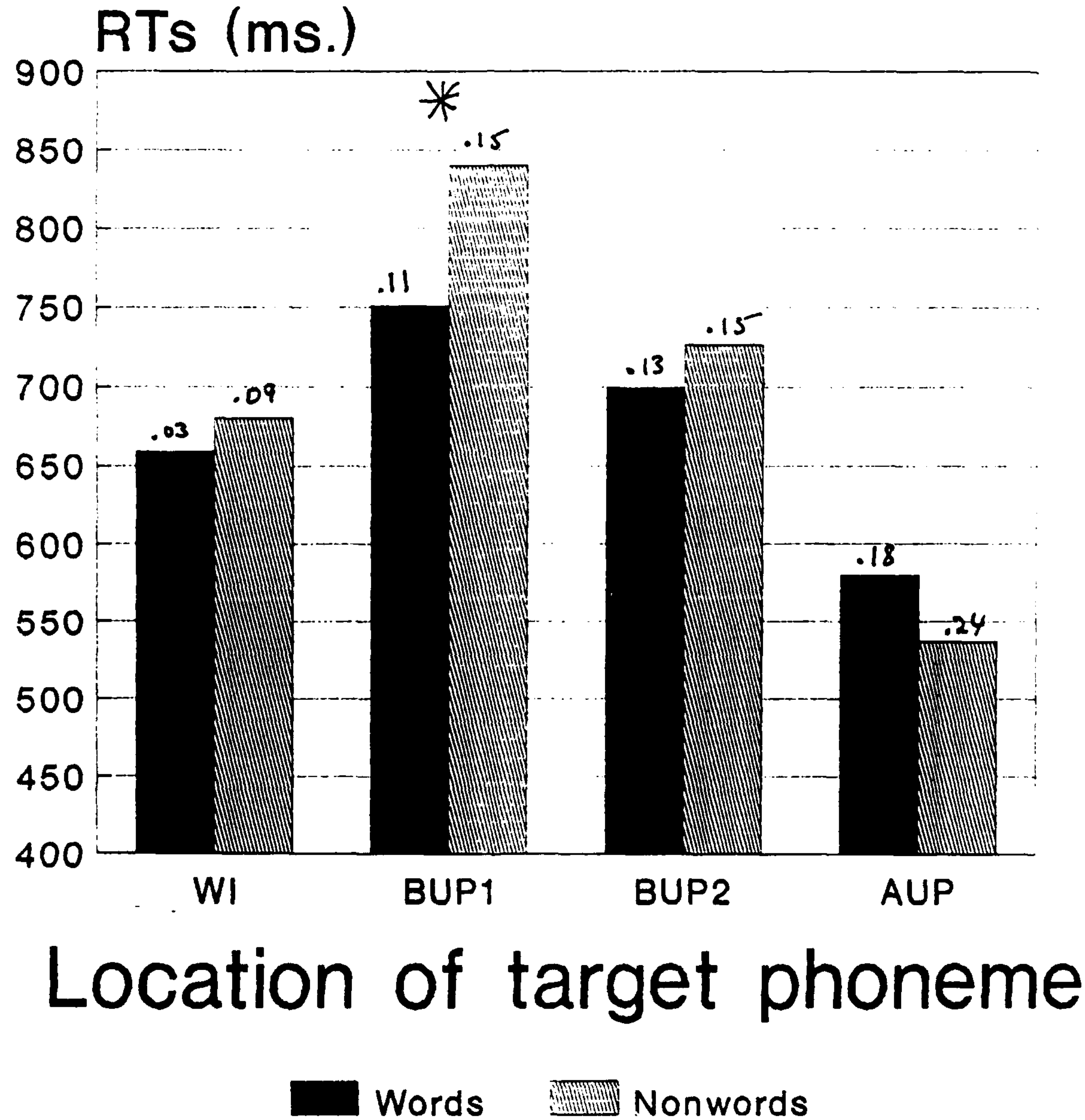
Early-Unique Words
S A R C A S T I C
(uniqueness point)
Reaction Time
900
850
800
750
700
650
600
550
500
450
400
(.03)
(.04)
(.13)
(.14)
WI
BUP1
Location of Target Phoneme
Words
Nonwords


Late-Unique Words
C I R C U M V E N T
(uniqueness point)
Reaction Time
900
850
800
750
700
650
600
550
500
450
400
(.06)
(.05)
(.09)
(.12)
(.15)
(.19)
WI
BUP1
BUP2
Location of Target Phoneme
Words
Nonwords

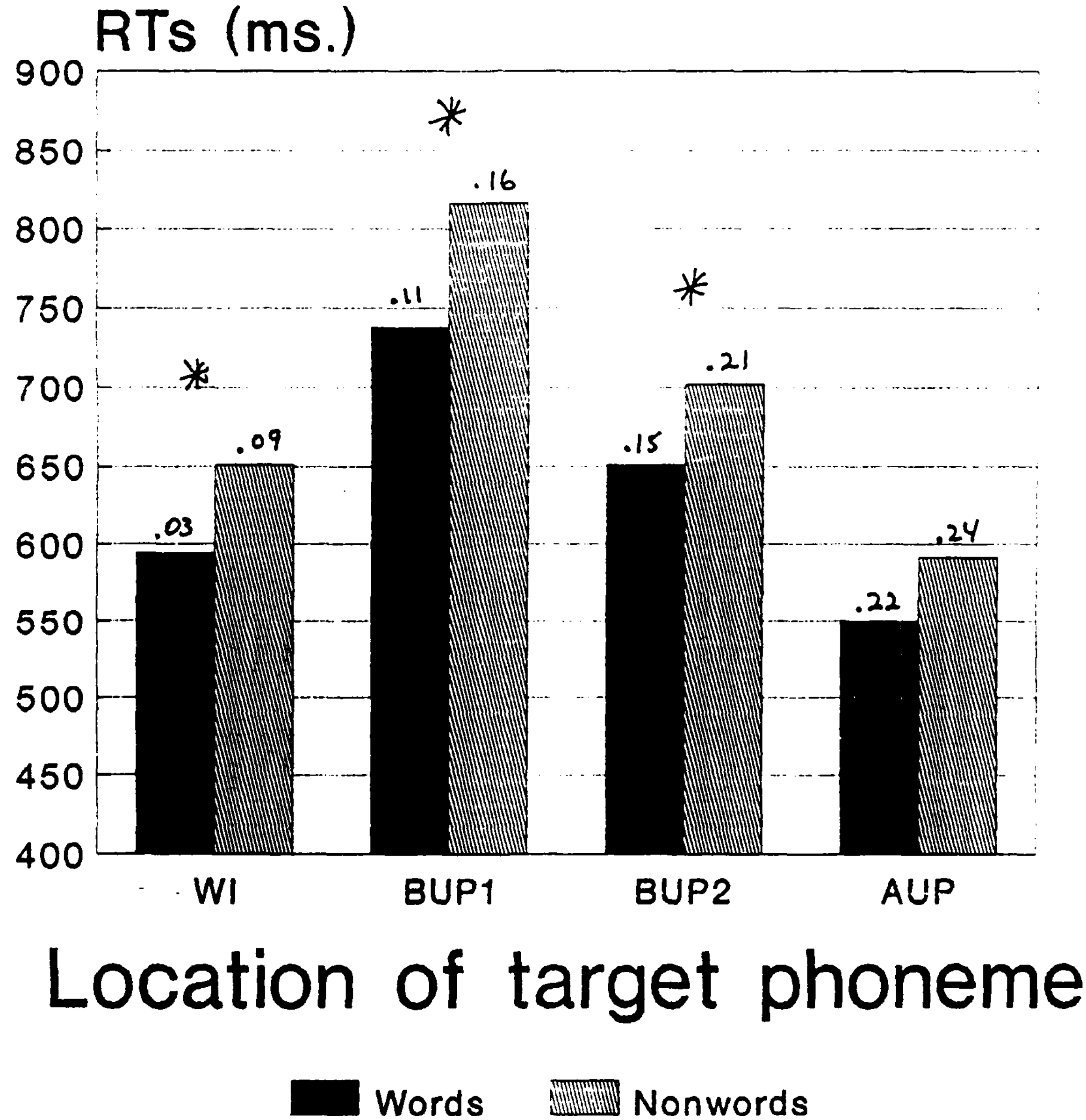
Early-Unique Words
Expansion experiment



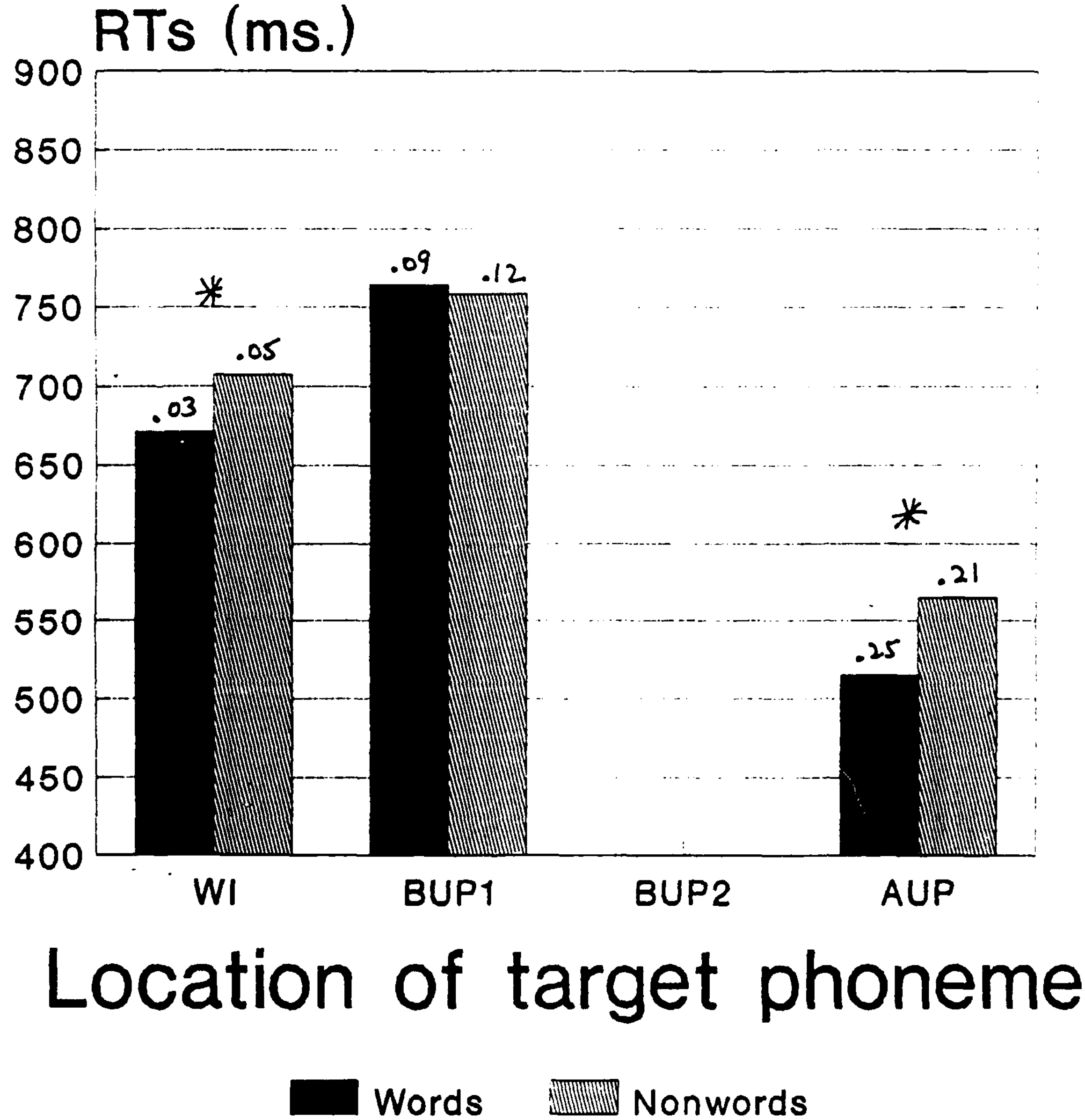
Late-Unique Words
Expansion exper (norm)



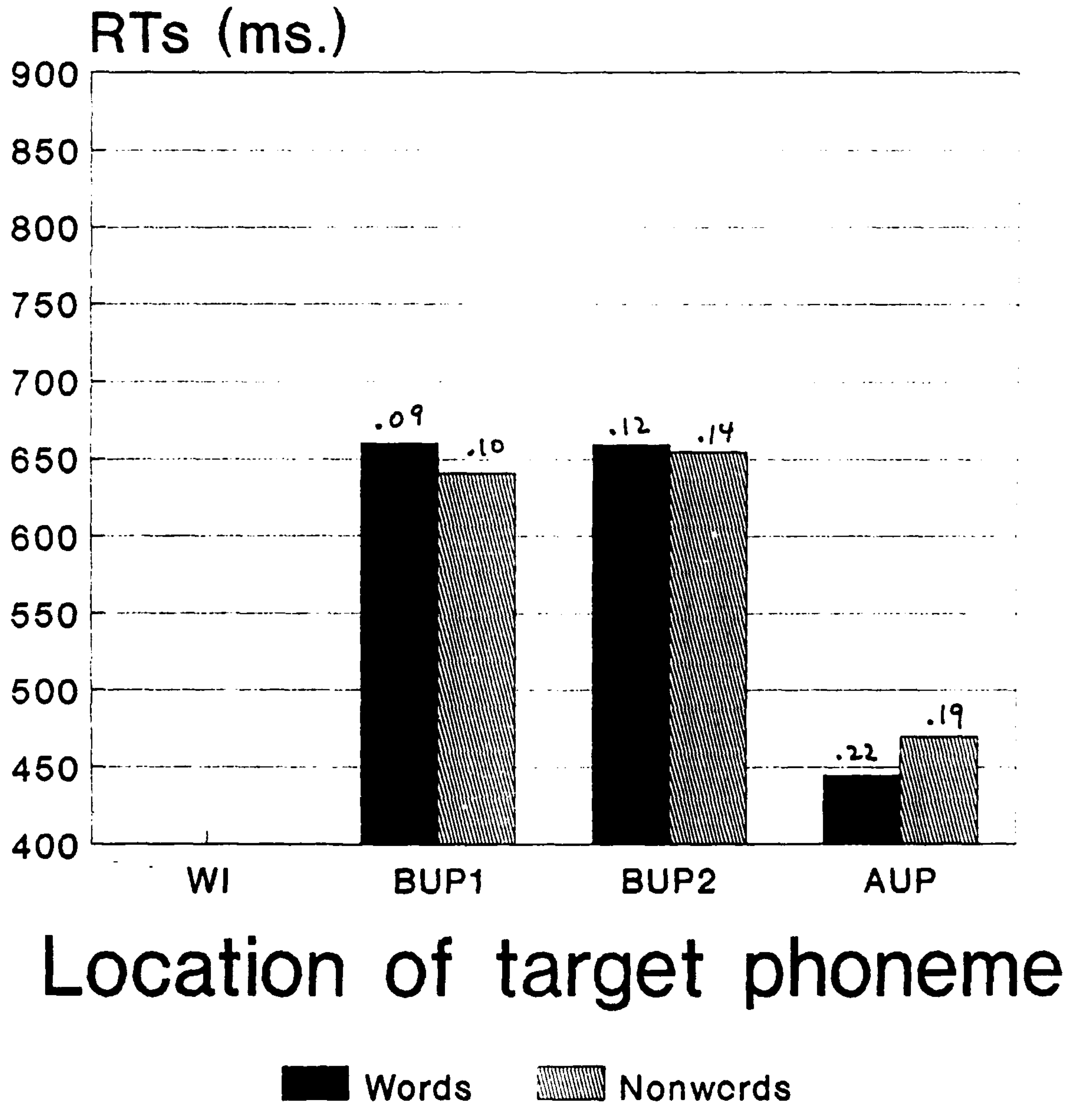
Late-Unique Words
Compression exper



Early-Unique Words
Compression exper (norm)



Late-Unique Parents
Beheading experiment



Early-Unique Parents

Beheading experiment